Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 428 (2008) 564–585
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Superoptimal approximation for unbounded symbols ☆

Claudio Estatico [a,b,*], Stefano Serra-Capizzano [c]

[a] Dipartimento di Matematica e Informatica, Via Ospedale 72, 09124 Cagliari, Italy
[b] Dipartimento di Matematica, Via Dodecaneso 35, 16146 Genova, Italy
[c] Dipartimento di Fisica e Matematica, Via Valleggio 11, 22100 Como, Italy

Received 28 September 2006; accepted 5 June 2007
Available online 16 June 2007
Submitted by E. Tyrtyshnikov

Abstract

The superoptimal Frobenius approximation of Toeplitz matrices is considered in connection with the case of unbounded symbols. In particular, we use the superoptimal approximation as preconditioner for the CG method when a Fisher–Hartwig singularity is present in the symbol, with special regard to systems coming from times series and financial applications. A theoretical discussion concerning classical circulant preconditioners and a numerical comparison with the Strang and with the optimal approximations are presented particularly with reference to the presence of noise.

AMS classification: 65F10; 15A30; 15A60; 47B48

Keywords: Toeplitz and circulant matrices; Preconditioning; Function and matrix approximation

1. Introduction

In this note we consider the superoptimal preconditioner proposed by Tyrtyshnikov [41] for the solution of Toeplitz systems. The superoptimal preconditioner is a special instance of

☆ This work was partially supported by MIUR, grant numbers 2004015437 and 2006017542.
* Corresponding author. Address: Dipartimento di Matematica e Informatica, Via Ospedale 72, 09124 Cagliari, Italy.
E-mail addresses: estatico@dima.unige.it (C. Estatico), stefano.serrac@uninsubria.it, serra@mail.dm.unipi.it (S. Serra-Capizzano).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2007.06.010

some approximate inverse preconditioners developed independently and used in the context of sparse solvers and, more generally, in the context of the convergence acceleration of Krylov type iterative methods (see e.g. [21,30,1,22] and references therein). The idea behind the superoptimal preconditioner is to approximate the system matrix by considering an optimization procedure concerning a kind of relative error in the matrix sense, instead of the absolute error considered in the classical optimal preconditioning [10]. As we will better discuss in the following, in principle the idea is very good and promising. Indeed, in the case of functions, it can be seen as the relative approximation has to be preferred to the absolute one: the former is able to give a good approximation also in parts of the domain where the function to be approximated shows small absolute values, whereas the latter may give wrong sign and in general is not able to match the position of the zeros. A matrix counterpart of these function approximation results can be found in [9,31], all related to the case of band-Toeplitz preconditioners for dense Toeplitz problems. However, in [15,7], by a direct eigenvalue analysis we proved that, surprisingly enough, the superoptimal approximation of the small eigenvalues is not good at all, since, basically, the small eigenvalues (if the matrix is positive definite) are replaced by eigenvalues of the order of the unity. A consequent presentation of this poor approximation of the small eigenvalues by Di Benedetto at the "Toeplitz Conference" in Cortona 2000 pushed Tyrtyshnikov to declare "today, I know that the superoptimal preconditioner is dead". On the other hand, in the same presentation a good side effect of this spectral results was presented. Indeed, if one consider a noisy problem and a Toeplitz matrix which has small eigenvalues in the noisy subspace, that is, in the space of the high frequencies, then this bad approximation is very useful: in fact, automatically, the superoptimal preconditioner acts as a regularizer that solves the problem in the space where the signal is concentrated (since therein its approximation of the system matrix is good) and ignores the noisy subspace (where the approximation is poor). The idea and the analysis have then been extended to the more involved case of image restoration problems, where the resulting structure is a two-level Toeplitz (or quasi-Toeplitz) matrix [12]. By an applicative point of view, the final message of this work is that the superoptimal approach is very convenient for moderate levels of noise and that it has to be combined with other techniques when the signal noise ratio (SNR) becomes of the order of the unity. However, it is important to recall that, in general, in the case where there is no noise or the problem is not ill-conditioned in the noisy subspace, it is not useful to use any regularization approach. Therefore, the complementary message is that, for solving an algebraic system of equations in the classical sense, that is, without any regularization, the superoptimal approach is not adequate and other preconditioning strategies have to be preferred at least in the following cases: when the original matrix has small eigenvalues, or when the original matrix is well conditioned. In the first situation, the superoptimal is definitely worse than other classical well-known preconditioning schemes [15,40]; in the second case it can have the same performance [7] (or slightly worse) but the computational cost of its more involving construction is superior, though of the same asymptotic order (see e.g. [12] and references therein).

The above mentioned results related to the behavior of the superoptimal approximation for ill-conditioned Toeplitz matrix have all been obtained in the context of Toeplitz sequences generated by bounded symbols with zeros, that is, in the context of Toeplitz systems such that the ill-conditioning is due to the zeros of the associated symbol (we notice that we have already considered and studied this case in [15,12], with special attention to signal/image restoration applications). Moreover, the proofs and the arguments of these works make use of mathematical tool which, besides non-trivial, often does not give a simple "intuitive" understanding. On these grounds, the contribution of this note is

(a) to discuss and explain the reasons behind the bad approximation of the small eigenvalues given by the superoptimal preconditioners, and
(b) to show that the only situation where the superoptimal approach can become competitive is when we add some noise (even in very small percentages i.e. close to the noise-free case).

With regard to the first item, we discuss here the relationships between the approximation of function and matrices with respect to relative distances, which represent the basis of the superoptimal approach, and we give a simpler meaning of its poor spectral approximation. With regard to the second item, here the novelty concerns the source of the ill-conditioning, which, differing from all the previous works, is due to a collective spectrum well separated from zero and asymptotically unbounded (i.e. the symbol is well separated from zero and unbounded, refer to [4] for a complete survey). We recall that such a situation is not just academic (for making "alive" a "dead technique"), but it comes from concrete applications in time series and financial problems (see the very interesting paper [23]). Moreover, although the paper is mainly devoted to (1D) Toeplitz systems, the results can be extended to multidimensional Toeplitz problems, as previously shown for bounded two-dimensional generating functions in [12].

The paper contains two further sections. In the next one we introduce the necessary definitions and we discuss the quality of the approximation for various well-known preconditioners, i.e. natural or Strang, optimal, and superoptimal. In the last we introduce examples of unbounded Toeplitz sequences coming applications, we perform numerical experiments (with and without noise), and we critically discuss the results: one of the main conclusions is that the case of Toeplitz systems with Fisher–Hartwig singularities considered in [23] is equivalent (up to a positive scaling and in a spectral asymptotic sense) to the case of symmetric blurring operators coming from classical signal/image processing.

2. Matrix and functions: a comparison between different approaches for approximation

We start with the definition of optimal, superoptimal, and natural (or Strang) preconditioners, with same basic arguments. Then we discuss their approximation properties, showing the relationships between (i) the relative and absolute approximations of the generating functions, and (ii) the corresponding approximations of the generated Toeplitz matrices. On these grounds, we will be able to understand the behavior of the superoptimal preconditioner for Toeplitz matrices with unbounded generating functions, as analyzed in the subsequent numerical section.

2.1. Optimal, superoptimal, and Strang preconditioners

Although the optimal and the superoptimal approaches can be considered in substantial more generality, here we restrict our attention to the Toeplitz case. Let f be a given Lebesgue integrable function on the interval $(-\pi, \pi)$ and let us consider its Fourier coefficients

$$a_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\mathrm{e}^{-\mathbf{i}j\theta}\,\mathrm{d}\theta, \quad \mathbf{i}^2 = -1, \quad j \in \mathbb{Z}. \tag{1}$$

From the function f (often called *generating function* or *symbol*), it is possible to construct the Toeplitz matrix $T_n(f)$ of size n where the entries of $T_n(f)$ along the jth diagonal coincide with jth Fourier coefficient a_j i.e.

$$(T_n(f))_{r,s} = a_{r-s}, \quad r, s = 1, \ldots, n. \tag{2}$$

If f is real valued then $a_j = \overline{a_{-j}}$ and therefore, for every k, the matrix $T_k(f)$ is Hermitian. Moreover, from well known results, the spectral behavior (asymptotic ill-conditioning, spectral localization and distributional results [32,33,43]) of these matrices and of their preconditioned versions is strongly characterized in terms of some analytical properties of the involved symbols. A subclass of Toeplitz matrices (to which it is not possible to attribute a symbol in the sense of the above definition) is the algebra $\mathscr{C}_n$ of circulant matrices: they share the algebraic property that every row is forward circular one-step shift of the previous row and where also the notion of previous has to be intended in a circular way. More precisely, the first row can be seen as the forward circular one-step shift of the last row. The latter nice algebraic feature translates to many properties related to circular convolutions. Here we only point out another important characterization in a spectral sense. Every circulant matrix of size n can diagonalized by the (unitary) discrete Fourier matrix. This means that A_n is circulant if and only if $A_n = F_n D F_n^*$ where D is complex diagonal

$$F_n = \left(\frac{1}{\sqrt{n}}\mathrm{e}^{-2\pi \mathbf{i} jk/n}\right), \quad k, j = 0, \ldots, n-1,$$

is the Fourier matrix and X^* denotes the complex transpose of X. Moreover, the diagonal matrix D has jth entry given by $p_n(\theta_j^{(n)})$ with $\theta_j^{(n)} = 2\pi j/n$, $j = 0, \ldots, n-1$, $p_n(z) = \sum_{k=0}^{n-1} a_k \mathrm{e}^{\mathbf{i}zk}$, $a_0, \ldots, a_{n-1}$ being the entry of the first column $c[1]$ of A_n. Notice that the above eigenvalue formula has also an important computational counterpart since the vector d containing the diagonal entries D is equal to $\sqrt{n}F_n^* c[1]$ so that the spectral decomposition of any circulant matrix can be recovered in $\mathrm{O}(n\log(n))$ complex operations via the celebrated FFT.

Definition 2.1. Let $\mathscr{C}_n$ be the algebra of circulant matrices and let $T_n(f)$ be a Toeplitz matrix associated to the symbol f. Then the following definitions hold:

- The Strang preconditioner (see [39]) $N_n(f)$ associated to $T_n(f)$ is the circulant matrix obtained from $T_n(f)$ by copying the first $[n/2]$ central diagonals with $[x]$ denoting the rounding of x. In other words, the jth entry of the first column and the jth entry of the first row of $N_n(f)$, $j = 0, \ldots, [n/2]-1$, are respectively the Fourier coefficients a_j and a_{-j} of f.
- The optimal preconditioner (see [10]) $C_n(f) = \mathrm{Opt}(T_n(f))$ is the unique solution of the minimization problem

$$\min_{X\in\mathscr{C}_n} \|A - X\|_{\mathrm{F}}, \quad A = T_n(f) \tag{3}$$

  with $\|\cdot\|_{\mathrm{F}}$ denoting the Frobenius norm i.e. the Euclidean norm of the vector which collects all the singular values or, equivalently, the Euclidean norm of n^2-sized vector obtained by merging in a unique vector all the columns of the argument.
- The superoptimal preconditioner (see [41]) $S_n(f) = \mathrm{Sopt}(T_n(f))$ is defined as the unique matrix whose Moore–Penrose pseudo-inverse $S_n^\dagger(f)$ solves the minimization problem

$$\min_{X\in\mathscr{C}_n} \|AX - I\|_{\mathrm{F}}, \quad A = T_n(f), \tag{4}$$

  I being the $n \times n$ identity matrix.

Some remarks are in order. The existence and uniqueness of the Strang, also called natural, preconditioner are implicit in the definition itself which clearly indicates an explicit expression. The existence and uniqueness of the optimal preconditioner (see [10]) follows from the strict convexity of the Frobenius norm that implies the existence and uniqueness of the minimizer from a given convex closed set. Indeed, since we are in a finite dimensional setting, clearly the linear space of the circulants $\mathscr{C}_n$ is closed and convex.

On this basis, the existence and uniqueness of the superoptimal operator (see [41]) comes from its explicit computations (see [7,15]) since

$$\mathrm{Sopt}^{\dagger}(A) = [\mathrm{Sopt}(A)]^{-1} = [\mathrm{Opt}(A^*A)]^{-1}(\mathrm{Opt}(A))^*, \tag{5}$$

whenever $\mathrm{Opt}(A^*A)$ is invertible e.g. when A is invertible. In addition, if A is also Hermitian and $\mathrm{Opt}(A)$ is nonsingular, it follows:

$$\mathrm{Sopt}(A) = [\mathrm{Opt}(A)]^{-1}\mathrm{Opt}(A^2) = \frac{\mathrm{Opt}(A^2)}{\mathrm{Opt}(A)}, \tag{6}$$

where the fractional notation emphasizes the commutativity of the matrix product inside $\mathscr{C}_n$. Therefore, the existence and uniqueness in the optimal case imply an analog statement in the superoptimal case. The case of singular A (which is less interesting in applications) can be treated similarly with a little more care. As final remark, we recall that the optimal approximation admits an easy to derive and very interesting representation since

$$\mathrm{Opt}(A) = F_n \mathrm{diag}(F_n^* A F_n) F_n^*, \tag{7}$$

where the operator diag applied to any square matrix X gives the diagonal matrix whose diagonal entries coincide with those of X.

In the next proposition we collect and discuss several spectral properties of these three approximations by focusing on the Toeplitz case with nonnegative generating functions (extensions are easily available but will complicate the notation without adding any further insight). Most of the results of the proposition can be found in the relevant literature, while others are briefly proven. It can be considered as a starting point, which will be accomplished in some sense by the following analysis concerning absolute and relative approximation in the next subsection.

As a prerequisite, we recall some notions and notations. We write that a function belongs to the Dini–Lipschitz class if its modulus of continuity $\omega_f(\cdot)$ is such that $\omega_f(\delta) = \mathrm{o}(1/|\log(\delta)|)$ for $\delta > 0$ (see e.g. [45]). Given a sequence $\{A_n\}$ of matrices with A_n of size n and a (Lebesgue) measurable function g defined over a set D equipped with finite and nonzero Lebesgue measure ($\mu(D) > 0$), we say that $\{A_n\}$ is distributed as g over D (in the sense of the eigenvalues) if for any continuous F with bounded support the following limit relation holds:

$$\lim_{n\to\infty} \Sigma(F, A_n) = \frac{1}{\mu(D)} \int_D F(g(\theta))\mathrm{d}\theta, \quad \Sigma(F, A_n) = \frac{1}{n}\sum_{j=1}^{n} F[\lambda_j(A_n)]. \tag{8}$$

Finally, a sequence $\{A_n\}$ (A_n of size n) is properly (or strongly) clustered at $s \in \mathbf{R}$ in the eigenvalue sense, if for any $\epsilon > 0$ the number of the eigenvalues of A_n not belonging to $(s-\epsilon, s+\epsilon)$ can be bounded by a pure constant q_ϵ possibly depending on ϵ but not on n. Here the term "properly (or strongly)" is replaced by "weakly" if q_ϵ is a possibly unbounded function of n with $q_\epsilon(n) = \mathrm{o}(n)$ $\left(\text{i.e.} \lim_{n\to\infty} \frac{q_\epsilon(n)}{n} = 0\right)$. We notice that $\{A_n\}$ is weakly clustered at s if and only if it is distributed as the constant function s.

In the following, if not explicitly mentioned, the norm $\|\cdot\|$ of a matrix is the spectral one.

Proposition 2.1. *Let f be a nonnegative L^1 function and, according to Definition 2.1, let us consider $N_n(f)$, $C_n(f)$, and $S_n(f)$ be the Strang, optimal, and superoptimal circulant approximations of $T_n(f)$. Then the following facts hold:*

1. *The Strang preconditioner $N_n(f)$ has eigenvalues $\mathscr{F}'_n[f](\theta_j^{(n)})$, $j = 0, \ldots, n-1$, where $n' = [n/2] - 1$, and $\mathscr{F}_q[f]$ is the Fourier sum of degree q of f (see e.g. [8]).*
2. *If f does not belong to the Dini–Lipschitz class but it is L^∞, then "anything" can happen (for instance, we can find f such that $N_n(f)$ is singular or indefinite even if $T_n(f)$ is positive definite, $N_n(f)$ is collectively unbounded even if $\|T_n(f)\| \leqslant \|f\|_\infty$ for every n). Furthermore, there exists f belonging to $L^1 \backslash L^2$ such that $\{N_n(f)\}$ is clustered at infinity even if $\{T_n(f)\}$ is distributed as f over $(-\pi, \pi)$.*
3. *If f has a finite number of zeros with maximal order $2k$ then the minimal eigenvalue of $T_n(f)$ is asymptotic to n^{-2k} [32] and the minimal eigenvalue of $N_n(f)$ is $\mathrm{O}(n^{-2k})$ if, in addition, f is smooth enough (see [42,13]). Moreover, if f is in the Dini–Lipschitz class and is 2π-periodic, then $\{T_n(f) - N_n(f)\}$ is properly clustered at zero.*
4. *If f belongs to the Dini–Lipschitz class and is 2π-periodic, then the eigenvalues of $N_n(f)$ converge to f in uniform norm on the Fourier grid sequence.*
5. *The optimal preconditioner $C_n(f) = \mathrm{Opt}(T_n(f))$ has eigenvalues $\mathscr{C}_{n-1}[f](\theta_j^{(n)})$, $j = 0, \ldots, n-1$, where $\mathscr{C}_q[f] = \frac{1}{q+1}\sum_{j=0}^{q} \mathscr{F}_j[f]$ is the Césaro sum of degree q of f (see e.g. [35]).*
6. *If f is L^∞, then $\|C_n(f)\| \leqslant \|T_n(f)\| \leqslant \|f\|_\infty$.*
7. *$\{T_n(f) - C_n(f)\}$ is weakly clustered at zero [36]. Moreover, if f is continuous and 2π-periodic, then the cluster is proper. Furthermore, if f has a finite number of zeros with maximal order $2k$, then the minimal eigenvalue of $C_n(f)$ is asymptotic to n^{-1} (see e.g. [11] for the case of a unique zero).*
8. *If f is continuous and 2π-periodic, then the eigenvalues of $C_n(f)$ converge to f in uniform norm on the Fourier grid sequence.*
9. *If f is not the zero function, then $T_n(f)$, $C_n(f)$, and $S_n(f)$ are positive definite for every n. For $N_n(f)$ this property is not guaranteed. If f is 2π-periodic, belongs to the Dini–Lipschitz class, and is strictly positive, then $N_n(f)$ is positive definite for n large enough.*
10. *The superoptimal preconditioner $S_n(f)$ is such that the eigenvalues are not infinitesimal in the related Fourier subspace where $T_n(f)$ and therefore, the optimal approximation show infinitesimal eigenvalues (see [15]).*
11. *The clustering properties of $\{T_n(f) - S_n(f)\}$ are of the same type as those of $\{T_n(f) - C_n(f)\}$ (see [15]).*

Proof. Items 1, 3, 5, 7, 10, 11 can be found in the relevant literature, as mentioned in the statements. Item 4 is a direct consequence of the fact that the Lebesgue constant of the Fourier sum is asymptotic (up to a multiplicative constant) to $\log(n)$ and therefore, the Fourier sum, whose uniform sampling gives exactly the eigenvalues of $N_n(f)$, has to converge to f since the modulus of continuity of f satisfies $\omega_f(1/n) = \mathrm{o}(1/\log(n))$ for every f in the Dini–Lipschitz class. Item 2 is a nice application of known facts. As first instance, the example of Du Bois–Raymond is a nonnegative function $f \in L^\infty$ with unbounded, highly oscillating Fourier sum (see e.g. [3]). Clearly the matrix $N_n(f)$ is unbounded and definitely indefinite while $T_n(f)$ is positive definite and uniformly bounded in spectral norm by $\|f\|_\infty$ (for the Toeplitz part see e.g. [38] where also the tools for proving Item 6 of Theorem 2.1 in [14] can be found). As second instance, for finding an example where $\{N_n(f)\}$ is clustered at infinity even if $\{T_n(f)\}$ is distributed as the symbol f over $(-\pi, \pi)$, it is enough to use the example of Kolmogorov (see e.g. [3]): the function belongs to L^1, but it is not in L^2 and has a Fourier sum diverging everywhere so that the eigenvalues of $N_n(f)$ collectively explode, but thanks to [43] it is still true that $\{T_n(f)\}$ is distributed as f over $(-\pi, \pi)$. Item 8 is trivial since (thanks e.g. to the beautiful theory by Korovkin) it is well known that the Césaro sum of any continuous function f, whose uniform sampling gives exactly the eigenvalues of $C_n(f)$, converges uniformly to f. For Item 6 by [38] we know that $\|T_n(f)\| \leqslant \|f\|_\infty$ whenever $f \in L^\infty$; in addition, since in general $\|\mathrm{Opt}(A)\| \leqslant \|A\|$ by (7), we have $\|C_n(f)\| \leqslant \|T_n(f)\|$ which proves the statement. Finally, concerning Item 9, the first statement related to $T_n(f)$, $C_n(f)$and $S_n(f)$ follows from (6) and (7), but the same result does not hold for $N_n(f)$, as already pointed out by Item 2; the last statement related to $N_n(f)$ follows from Item 4, since now f is strictly positive. □

Remark 2.1. The well known Item 1 in the above proposition has an interesting consequence. Take $f \in L^\infty$ and consider $N_n(f)$. Since the entries of $N_n(f)$ contain exactly the same coefficients as $T'_n(f)$ with every Fourier coefficient counted $2n'$ times, thanks to the Parseval equality it follows that $\|N_n(f)\|_{\mathrm{F}}^2 = 2n'\|\mathscr{F}'_n[f]\|_{L^2}^2$. Thus, from $2n'\|\mathscr{F}'_n[f]\|_{L^2}^2 \leqslant n\|f\|_{L^2}^2 \leqslant n\|f\|_\infty^2$, with $\|h\|_{L^2}^2 = \frac{1}{2\pi}\int_{[-\pi,\pi]} |h|^2$, by virtue of the spectral decomposition of $N_n(f)$ in Item 1, we have

$$\|N_n(f)\|_{\mathrm{F}}^2 = \sum_{j=0}^{n-1} |\mathscr{F}'_n[f](\theta_j^{(n)})|^2 \leqslant n\|f\|_\infty^2.$$

Consequently, the cardinality of the set of indices j such that $\mathscr{F}'_n[f](\theta_j^{(n)})$ is unbounded as n tends to infinity has to be $\mathrm{o}(n)$. This means that the set of grid points in which the Fourier sum can diverge is negligible and more precisely its cardinality is $\mathrm{o}(n)$. Taking into account the possible maximal growth of a polynomial of degree $n' = [n/2] - 1$, it follows that the set where the Fourier sum can diverge in $[-\pi, \pi]$ has to be of zero Lebesgue measure and this is a linear algebra view of a Carlesson-type result (see e.g. [3]).

2.2. *Matrix and functions: relative and absolute approximations*

As observed in the previous subsection, in many contexts and for a theoretical analysis of Toeplitz preconditioning performances, a fundamental tool is the use of the generating function and of its analytic properties. Therefore, we consider two types of approximation and two types of problem, the first referring to functions and the second to matrices. To be precise, we should talk of matrix sequences instead on matrices since, as it was also clear from many items of Proposition 2.1, the natural setting is the one of sequences of matrices of increasing dimension, where we are interested in limit properties. This fact has, of course, an applicative counterpart in the solution of large numerical problems.

Problem 1. Given $f \in C_{2\pi}$ (space of continuous 2π-periodic functions) with $f \geqslant 0$ vanishing at most in a finite number of values in $[0, 2\pi)$, find $g \in \mathscr{P}_s$ (space of 2π periodic trigonometric polynomials of degree at most s) such that $E(f, g)$ is minimized ($E(\cdot, \cdot)$ error function to be specified later).

Problem 2. Given $T = \{T_n(f)\}$, sequence of Toeplitz matrices with f as in Problem 1, find $P = \{P_n\}$ ($P_n \in \mathscr{S}_n$, where $\mathscr{S}_n$ denotes a fixed space of "simpler" matrices) sequence of approximations such that $E_n(T, P)$ is minimized ($E_n(\cdot,\cdot)$ error function to be specified later) at least for large n.

Before going on, we can summarize that, with respect to the second problem, two main approaches have been followed by the Toeplitz scientific community: the first is related to $\mathscr{S}_n = \mathscr{C}_n$ space of circulant matrices (or any other convenient (trigonometric, wavelet, etc.) algebra related to unitary fast transforms), the second is $\mathscr{S}_n = \mathscr{B}_n(s)$ space of s-band Toeplitz matrices of size n, that is, Toeplitz matrices generated by trigonometric polynomials in $\mathscr{P}_s$.

Now we can go further and consider two types of approximations dictated by the choice of error functions E (Problem 1) and E_n (Problem 2).

Dealing with Problem 1, for function we can consider absolute and relative approximations, that is, $E(f, g)$ can be one of the following:

$$E_a(f, g) = \|f - g\|, \tag{9}$$

$$E_r(f, g) = \|f/g - 1\| \tag{10}$$

with $\|\cdot\|$ suitable norm: here we restrict the attention to $\|\cdot\| = \|\cdot\|_\infty$ (the sup norm), although the L^2 norm could be of interest too.

The first approximation, that is, the absolute one, of the nonnegative function f leads to essential problems, since the nonnegativity of the minimizer is not guaranteed and, to impose nonnegativity, the quality of the approximation will be partly spoiled. Moreover, also the matching of the zeros is of course not guaranteed [24]. On the other hand, the nice implication of the second approach, that is, the relative one, is that, in presence of zeros of even orders (which is automatically true if f is nonnegative and smooth enough), the nonnegativity is for free and also a perfect matching to the zeros is guaranteed by the minimizer if the degree s of the trigonometric polynomial is large enough (see [31] and references therein). In particular, by choosing the infinity norm, if we set

$$\min_{g \in \mathscr{P}_s} \|f/g - 1\|_\infty = r_s^*,$$

then r_s^* tends to zeros as s tends to infinity and this convergence is fast, e.g. exponentially fast if $h = f/g_{\min}$ is infinitely differentiable. Here $g_{\min}$ is the nonnegative polynomial of minimal degree such that $f/g_{\min}$ is continuous and strictly positive. Moreover, the minimizer g_s^* exists and it is unique, it is nonnegative if $s \geqslant \text{degree}(g_{\min}) = s_{\min}$, and, in that case, is such that $g_s^* = g_{\min} \cdot g_{s-s_{\min}}$ with $g_{s-s_{\min}}$ nonnegative polynomial.

It is interesting to notice that the relative function approximation (10) have already been used to compute circulant preconditioners. In particular, Chan and Tang [9] and the second author [31] have found excellent preconditioners where the spectral equivalence is guaranteed and it is also possible to have a proper clustering without outliers (see [31]). The reason of this good spectral approximation basically relies on the fact that the Toeplitz operator from L^1 to n-by-n matrices is a linear positive operator which implies that the eigenvalues of $T_n^{-1}(g)T_n(f)$ are in the interval $(\inf f/g, \sup f/g)$ if $g \geqslant 0$, g not identically zero, and with $\inf f/g < \sup f/g$. Therefore, in the case of our relative minimization process, we have

$$\|T_n^{-1/2}(g)T_n(f)T_n^{-1/2}(g) - I\| < r_s^*$$

with respect to the spectral norm. Notice that in the case of exponential or polynomial decay to zero of s, the minimization in infinity norm of $f/g - 1$ (which is done by a modified Remez algorithm) can be replaced by a more convenient technique by maintaining (up to, at most, a $\log(s)$

factor) the same exponential or polynomial approximation. To this aim, it is enough to consider a quasi-optimal approximation technique: for instance, we can replace g_s^* by $\tilde{g}_s^* = g_{\min} \cdot \tilde{g}_{s-s_{\min}}$ where $\tilde{g}_t$ is the Chebyshev interpolation of $f/g_{\min}$ of degree t and $g_{\min}$ is numerically determined as in [34,37].

Conversely, the use of $T_n(g)$ (with g minimizer with respect to the absolute error (9)) can be quite disappointing because we can have negative eigenvalues in the preconditioner which is not even guaranteed to be invertible.

We now switch to Problem 2, and consider also in this case two kinds of approximation which can be used as matrix error function E_n, that is

$$E_{n,a}(T, P) = \|T_n(f) - P_n\|, \tag{11}$$

$$E_{n,r}(T, P) = \|T_n(f)P_n^{\dagger} - I\|, \tag{12}$$

where $T = \{T_n(f)\}$, $P = \{P_n\}$, and the chosen norms are the spectral norm (maximal singular value) and the Frobenius norm. From a computational viewpoint, the spectral norm is really difficult to handle while the Frobenius norm (which comes from a positive scalar products and makes the space of the matrices an Hilbert space) is convenient for both practical and theoretical derivations.

Now if one consider $P_n \in \mathscr{B}_n(s)$ and one tries to solve the two problems (11) and (12), by a direct use of asymptotic results in [33], it follows that the solution is asymptotically close (for large n) to $T_n(g)$ where g solves the corresponding function problem (9) or (10). Therefore, by virtue of the previous arguments related to the function approximations, this tells us automatically that the relative error approach is definitely better and has to be preferred. This conclusion is of course true in the band Toeplitz case, but we should refrain from claiming this fact as a general fact. In this respect, we remark that a counterexample has been already analyzed and corresponds to the approximation in the classical algebra of circulant matrices, that is, when $\mathscr{S}_n = \mathscr{C}_n$. In particular, when looking at Proposition 2.1, we find what seems a philosophical contradiction: while for functions f the relative error approach is the only way for catching the exact location and the exact order of the zeros, for Toeplitz matrices $T_n(f)$ this is totally false. Indeed, the superoptimal approach $S_n(f)$, which comes from a relative approximation, gives wrong results just for the smallest eigenvalues, since the eigenvalues which are asymptotically infinitesimal in $T_n(f)$ are replaced by eigenvalues in $S_n(f)$ whose liminf is strictly positive as the size tends to infinity (notice that in the same situation the eigenvalues of the optimal preconditioner $C_n(f)$ tend to zero).

Of course, as already mentioned in the Introduction, if one looks at the mathematical proof in [15] then one will understand the reason of this different spectral distribution between $T_n(f)$ and $S_n(f)$. But this understanding is a *technical* understanding and we would like to explain the real reason behind this unnatural behavior. To do this, we limit ourselves to the case where f is a nonnegative, nonzero trigonometric polynomial with at least one zero, and we look closely to the minimization process involved in the determination of the superoptimal approximation, by considering the two following facts:

Fact 1. $T_n(f) = N_n(f) + R_n(f)$ where $N_n(f)$ is the Strang approximation and $R_n(f)$ has rank equal to $2r$ with r degree of f. Moreover, as n tends to infinity the spectral norm of $R_n(f)$ has strictly positive liminf and the eigenvalues of $N_n(f)$ are the sampling of f on the circulant grid points (in this case it is clear that the Fourier sum equals the function itself for n larger than the degree).

Fact 2. Looking at the matrix whose norm has to be minimized, we can write $T_n(f)P_n^\dagger - I = X_n(1) + X_n(2) - I$, where $X_n(1) = N_n(f)P_n^\dagger$ and $X_n(2) = R_n(f)P_n^\dagger$. Moreover, $R_n(f)$ is a $2r$ rank symmetric matrix with null diagonal. Therefore, any invertible principal minor M of $R_n(f)$ has even rank $2, 4, \ldots, 2r$ with null diagonal. From an elementary interlacing argument, it follows that M has j negative eigenvalues and j positive eigenvalues whenever its rank is $2j$, $j = 1, \ldots, r$. As a consequence the global correction matrix $R_n(f)$ has r positive eigenvalues and r negative eigenvalues and, by similarity, the same is true for $X_n(2)$, if P_n is invertible. In the case where P_n is singular, due to the symmetry inherited by P_n, it is evident that the rank of $X_n(2)$ is $2j$ for a given $j \in \{1, \ldots, r\}$ and again with exactly j positive and j negative eigenvalues.

Now we can move on a sort of "reductio ad absurdum" to explain, as mentioned, the real reason behind the bad spectral approximation given by the superoptimal preconditioner.

First notice that, if P_n "imitates" the original function f, then it should have minimal eigenvalues going to zero as n^{-2k} with $k \geqslant 1$. In this case the term $X_n(1)$ is bounded and nonnegative definite (possibly equal to the identity matrix), but the term $X_n(2) = R_n(f)P_n^\dagger$ will have eigenvalues exploding to ∞ (see the Rayleigh quotient arguments in [29]). As a result, since $X_n(1)$ is nonnegative definite and $X_n(2)$ is similar to a Hermitian matrix, the sum $X_n(1) + X_n(2)$ will have spectral norm exploding to infinity and therefore, $T_n(f)P_n^\dagger - I$ will be unbounded in spectral norm and, a fortiori, in Frobenius norm as n tends to infinity. Thus the superoptimal preconditioner, which is the minimizer of $\|T_n(f)X^\dagger - I\|_{\mathrm{F}}$ among all the circulant matrices X, cannot mimic the minimal eigenvalues of T_n.

Therefore, we have understood which is the problem, all summarized in the last remarks. Indeed, although the relative approximation in principle is exactly what is needed for better approximating a function where it is small, in this case of circulant approximation to a Toeplitz structure we have a rank obstruction. Basically, the small rank term $R_n(f)$ has the role of becoming an unsurmountable obstacle to the possibility of the minimizer P_n to have small eigenvalues in the subspaces where the original operator $T_n(f)$ has small eigenvalues.

Remark 2.2. Given $T_n(f)$ with nonnegative f and with a finite number of zeros and of finite orders, we should acknowledge that the above latter discussion was also the main idea for proving that circulant preconditioners cannot be spectrally equivalent to $T_n(f)$ (uniformly with respect to n) in the single-level case and cannot insure essential spectral equivalence (i.e. up to a constant number of unbounded outlying eigenvalues) in the multilevel case (see [29] and references therein). We also observe that the impossibility of the plain spectral equivalence was already contained (in a slightly weaker sense) in [28], where Manteuffel and Parter used infinite dimensional arguments from a PDE context.

3. Unbounded symbols, classical and regularizing approximations

In the present section, we give some numerical evidences on how the distribution of the eigenvalues of the superoptimal approximation shown in Section 2 can be favorable for solving ill-conditioned Toeplitz systems with unbounded symbols. We are interested in the solution of linear systems when the right-hand side is corrupted by noise and in the classical case where the noise is not present. In the noisy case, the preconditioned conjugate gradient least squares (PCGLS) method is used as a regularizer, i.e. we are not interested in computing the solution of the algebraic system but we stop the iterations when the computed solution has minimal distance with respect to the true solution. Here for true solution we mean the solution of the linear system related to the exact right-hand side, i.e. with zero noise. Of course, the previous stopping criterion is not practical and realistic since the true solution is unknown but there exist quite good techniques for evaluating the optimal iteration (see e.g. [17,6]). Finally, we recall that in this regularization process it is crucial to neglect those components which carry noise information. In this respect, the use of the superoptimal approximation as regularizing or filtering preconditioner has been proposed in [15] and analyzed in [12,18,19] with regard to inverse problems arising in signal/image restoration.

In our tests, we first consider the $n \times n$ Toeplitz matrix $T_n(f)$ generated by the function

$$f(\theta) = \frac{1}{2\pi}\sigma_s^2|2\sin(\theta/2)|^{-2d}, \tag{13}$$

where $\sigma_s = 0.27^{1/2}$ and $d \in (0, 1/2)$ are fixed real parameters, as prototype of Toeplitz matrices arising in Time Series analysis, according to Brockwell and Davis [5]. Notice that the Toeplitz matrix $T_n(f)$ has a Fisher–Hartwig singularity due to the pole of order $2d$ at zero [23]. Thanks to the explicit computation of the Fourier coefficients of f, we have $(T_n(f))_{r,s} = a_{r-s}$ where

$$a_j = \frac{\sigma_s^2\Gamma(1-2d)\Gamma(j+d)}{\Gamma(d)\Gamma(1-d)\Gamma(j-d+1)} \quad j = 0, 1, \ldots, n-1 \tag{14}$$

with $\Gamma(x) = \int s^{x-1}\exp(-s)\mathrm{d}s$ being the Gamma function. By a computational point of view, to avoid overflow for the terms $\Gamma(j+d)$ and $\Gamma(j-d+1)$ for large values of j, the logarithm L_Γ of the Gamma function is considered as follows $a_j = \sigma_s^2\exp(L_\Gamma(1-2d) + L_\Gamma(j+d) - L_\Gamma(d) - L_\Gamma(1-d) - L_\Gamma(j-d+1))$.

In addition, since we are also interested in a (very) high ill-conditioning, we consider the Toeplitz matrices associated with f^2 and f^3, by computing the appropriate convolution of the vector $(a_{-2n}, a_{-2n+1}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{2n-1}, a_{2n})$ with itself. We remark that $T_n(f)$ with f as in (13), $d \in (0, 1/2)$, is well defined in the sense of (2) since the symbol f belongs to L^1 over $(-\pi, \pi]$ and therefore, its Fourier coefficients exist. On the other hand, for $c = 2, 3$ we are in general not allowed to write $T_n(f^c)$, since f^c fails to belong to L^1 over $(-\pi, \pi]$ if $2cd \geqslant 1$. Therefore, we consider the new operator $\widetilde{T}_n(f^c)$ where $\widetilde{T}_n(f) = T_n(f)$ and $\widetilde{T}_n(f^c)$, $c = 2, 3$, is defined through c convolutions of the Fourier vector of f.

The true solution vector $x = (x_1, x_2, \ldots, x_n)^t$ is the sum of two different "impulses", $x_j = 0.5k_{0.1}(p_j + 0.9) + k_{0.05}(p_j - 0.8)$, where $k_\sigma(t)$ denotes the Gaussian distribution with zero mean and standard deviation σ and the points p_j for $j = 1, 2, \ldots, n$ are uniformly spaced in $[-2, 2]$ (see [16] for details).

The right-hand side vector b is the sum of the "true" object $T_n x$ and Gaussian white noise, that is, $b = T_n x + \eta$, where η comes from a normal distribution with zero mean. Several levels of noise have been tested, corresponding to different values of the relative error $\|\eta\|/\|T_n x\|$.

The numerical code have been implemented on IDL 5.4 (Interactive Data Language) with floating-point precision of about 10^{-16}.

The following table summarizes the numerical tests:

Toeplitz matrices $T_n = \widetilde{T}_n(h)$:

I *Test FH1*

h is the unbounded function f whose expression is given in (13), with Fisher–Hartwig singularity of order $2d$.

II *Test FH2*

h is the unbounded function f^2, with Fisher–Hartwig singularity of order $4d$.

III *Test FH3*

h is the unbounded function f^3, with Fisher–Hartwig singularity of order $6d$.

Poles: for each one of the three tests FH1, FH2 and FH3, we consider three coefficients $d = 0.37, 0.49, 0.499$.

Matrix dimensions: for each one of the three tests FH1, FH2 and FH3, we consider three instances $n = 256, 512, 1024$.

We observe that the conditioning of $T_n = \widetilde{T}_n(h)$, $h = f^c$, f as in (13), grows asymptotically as n^s where $s = 2cd$ is the order of the Fisher–Hartwig singularity.

3.1. The spectral distribution

Table 1 shows the "absolute" Frobenius distance $\|T_n - P_n\|_F$ and the "relative" Frobenius distance $\|T_n P_n^\dagger - I\|_F$ for the natural, optimal, and superoptimal preconditioners in the algebra of circulant matrices. We show as instance the results related to $n = 1024$, and we remark that the behavior is very similar for all the considered tests. According to the theory, the T. Chan optimal preconditioner minimizes the first distance, whereas the Tyrtyshnikov superoptimal preconditioner minimizes the second one (really, there is one case which does not satisfy this rule – see the TEST FH3 for $d = 0.499$ on the left side of the latter row – where the value of the optimal preconditioner is larger than the value of the natural one, which, in our opinion, is only due to high numerical instabilities since the norm of the matrix T_n is about 10^{20}).

What is really interesting to gather from Table 1 is that the superoptimal preconditioner leads to very different values with respect to the others two preconditioners when the "unboundedness" (and hence the ill-conditioning) becomes large.

Table 1
Frobenius distances related to the natural, optimal, and superoptimal preconditioners in the algebra of circulant matrices, for $n = 1024$

d	$\|T_n - P_n\|_F$			$\|T_n P_n^\dagger - I\|_F$		
	Natural	Optimal	SuperOpt	Natural	Optimal	SuperOpt
TEST FH1						
0.37	8.3747	7.7225	7.9473	4.3828	4.3580	4.3087
0.49	39.377	35.924	38.457	16.279	16.181	14.439
0.499	44.186	40.284	43.805	18.084	17.972	15.671
TEST FH2						
0.37	250.83	225.80	387.98	362.54	349.56	31.655
0.49	5318.7	4757.6	65890	19441	5350.6	31.944
0.499	6700.5	5991.8	1.39e08	57776	6172.4	31.980
TEST FH3						
0.37	7300.4	6518.4	149601	14254	10894	31.952
0.49	776,140	691,667	3.887e10	12,422	14,172	31.993
0.499	1106935	1256211	1.909e16	12413	6529.7	32.039

In the first case (TEST FH1), related to the lowest ill-conditioning, the absolute distances $\|T_n - P_n\|_F$ are similar for the three considered preconditioners, and the same arises for the relative distances $\|T_n P_n^{\dagger} - I\|_F$, regardless the values of the parameter d and the dimension n (see the top of table). However, this is not true in general, and indeed the scenario changes deeply in the other two tests (TEST FH2 and TEST FH3) related to higher ill-conditioning. Now (see the middle and the bottom of the table) the values are very dissimilar, and this shows that the superoptimal preconditioner is "strongly" different from the other two ones, according to the arguments of the previous section. As an example, in the TEST FH2 with $d = 0.499$, the optimal preconditioner gives rise to the absolute distances $\|T_n - P_n\|_F$ of about 6000, whereas the superoptimal preconditioners leads to the value of about 1.4×10^8. In the same case, the optimal preconditioner implies relative distances $\|T_n P_n^{\dagger} - I\|_F$ of about 6200, whereas the superoptimal preconditioners leads to values of about 32. This behavior is amplified in the third case of TEST FH3. To summarize, the analysis of this first table confirms that the minimizations behind the optimal and of superoptimal preconditioners act in different manners, leading to completely different circulant approximations (at least in some critical subspaces) when the ill-conditioning is high.

Another aspect which is interesting to study concerns the condition numbers. Table 2 shows the condition numbers of the Toeplitz matrices T_n, the natural N_n, optimal C_n and superoptimal S_n preconditioners, and the natural $N_n^{\dagger} T_n$, optimal $C_n^{\dagger} T_n$ and superoptimal $S_n^{\dagger} T_n$ preconditioned matrices in the same setting of Table 1, i.e. $n = 1024$. As expected, although the natural and optimal behave similarly, again the superoptimal exhibits very different results.

Here, the main fact to remark is that the condition numbers of the superoptimal preconditioner stay well bounded, even and especially when the Toeplitz matrix becomes very ill-conditioned (see the left side of Table 2). From the latter table, we can infer that the superoptimal S_n is not able to approximate the spectrum of T_n, and that this property is amplified when the Toeplitz systems become very difficult to manage. As an example, we can consider the TEST FH2 with $d = 0.49$: we have that the condition number of the Toeplitz matrix, $K_2(T_n)$, is about 3.1×10^9, $K_2(N_n)$ and $K_2(C_n)$ are respectively about 9.4×10^9 and 1.8×10^9, while the condition number of the superoptimal preconditioner $K_2(S_n)$ is appreciably lower, about 5.6×10^4. The same

Table 2
Spectral condition numbers of the natural, optimal, and superoptimal preconditioners and preconditioned matrices, for $n = 1024$

d	$K_2(T_n)$	$K_2(N_n)$	$K_2(C_n)$	$K_2(S_n)$	$K_2(N_n^{\dagger} T_n)$	$K_2(C_n^{\dagger} T_n)$	$K_2(S_n^{\dagger} T_n)$
TEST FH1							
0.37	643.1	643.3	616.2	608.0	510.2	516.9	534.5
0.49	36,661	36,694	36,408	28,988	5004	5184	4732
0.499	409,900	410,294	408,379	309,377	5947	6173	5437
TEST FH2							
0.37	498,448	510,704	477,831	3456	624,758	664,513	17,402
0.49	3.123e09	9.404e09	1.867e09	56738	1.345e08	5.355e07	354084
0.499	4.120e11	2.931e12	2.221e11	3756	4.109e08	7.016e07	3.123e08
TEST FH3							
0.37	5.528e08	1.777e08	1.130e08	1114.6	2.679e08	2.664e08	2550018
0.49	5.292e19	1.231e14	1.143e14	8.309	9.663e09	4.4489e11	7.473e15
0.499	7.723e19	5.003e14	3.245e14	4.823	6.348e12	6.903e11	9.952e14

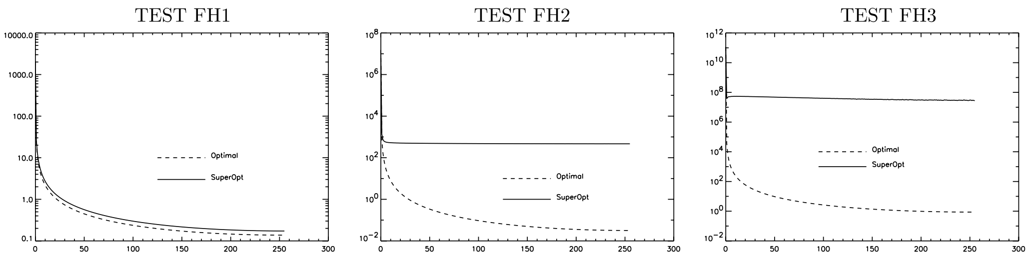

Fig. 1. Spectral distribution of optimal versus superoptimal preconditioers, for $d = 0.49$ and $n = 512$.

results (in a stronger version) can be observed in all the subsequent cases of TEST FH3 (see for instance the latter case at the bottom of the table, where $K_2(T_n)$, is of the order of 10^{19}, while $K_2(S_n)$ is of the order of unity!). Fig. 1 shows the spectral distribution of the optimal and superoptimal preconditioners, for $n = 512$, and $d = 0.49$. We plot only the eigenvalues related to positive Fourier frequencies, since the spectrum is symmetric. The graphs show that the highest eigenvalues, related to the pole in the origin, are well approximated by both the preconditioners. On the contrary, the higher is the degree of the pole, and hence the ill-conditioning, the worse is the approximation by the superoptimal in the rest of the spectrum.

3.2. *The regularization effectiveness*

All the next tables show the best Relative Restoration Error (RRE) $\|x_k - x\|/\|x\|$ in the Euclidean norm, and the corresponding iteration k obtained by all the preconditioners, where x denotes the true solution and x_k the solution at the kth iteration of the (P)CGLS.

It is important to mention here a result of Kailath and Olshevsky [25], stating that in principle one iteration of PCGLS with a trigonometric transform based preconditioner can be implemented at the same cost as a (non-preconditioned) CGLS iteration (we did not apply this trick in our experiments); therefore, the number of iterations performed is a good parameter for evaluating and comparing the performances of the preconditioners.

We remark that the superoptimal preconditioner provides a “fixed” level of regularization, i.e. of filtering, depending only on the size n and on the generating function f. This means that for some problems, severely ill-posed or with high noise, it could be necessary to strengthen the regularization properties, while in other circumstances, this filtering could be too high and useless, leading to a slower convergence speed with respect to the optimal one. In our tests, we have found examples of all these different scenarios. However, in presence of such difficulties, the “level” of regularization can be controlled by using a parameter-dependent family of generalized superoptimal preconditioners developed in [18], where the additional parameter plays the role of regularization parameter.

3.2.1. *TEST FH1*

As first instance, we can say that the ill-conditioning related to the first test FH1 is not too high to require high regularization features. In this case, the natural and the optimal preconditioners are better than the superoptimal one, since the RREs are similar, and the formers require less iterations. Table 3 reports only the case with $d = 0.37$, and $n = 1024$, but the behavior is similar for all the other instances with the different parameters d and n. Fig. 2 shows the corresponding convergence history in the first 100 iterations for the noiseless case.

Table 3
Best relative restoration errors (RREs) and number of iterations for TEST FH1

	TEST FH1: $d = 0.37, n = 1024$						
Noise	0	5.00e−09	5.00e−08	5.00e−07	5.00e−06	5.00e−05	5.00e−04
No prec	1.3311e−13	1.0214e−06	1.0192e−05	9.4034e−05	6.0103e−04	4.2278e−03	2.8709e−02
	96	41	32	20	13	10	7
Nat	1.1911e−13	1.0216e−06	1.0216e−05	1.0216e−04	1.0216e−03	1.0216e−02	1.0216e−01
	14	17	17	16	15	14	15
Opt	1.2488e−13	1.0216e−06	1.0216e−05	1.0216e−04	1.0216e−03	1.0216e−02	1.0216e−01
	14	13	13	10	10	7	5
Super	1.2769e−13	1.0216e−06	1.0216e−05	1.0216e−04	1.0216e−03	1.0216e−02	1.0213e−01
	20	13	13	13	11	8	5

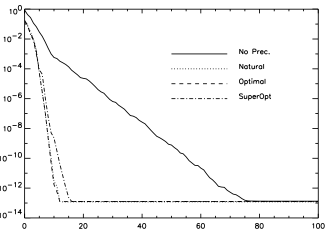

Fig. 2. RREs versus Iterations for TEST FH1 ($d = 0.37$, $n = 1024$), noiseless input data, within the first 100 iterations.

The conclusion is that the three preconditioners act similarly in terms of reconstruction quality, and in particular, the superoptimal is slightly worse in terms of convergence speed. This behavior confirms some recent convergence results just related to this test case, developed first by Hurvich and Lu [23] for the optimal preconditioner, and extended by Vong et al. for [44] to the superoptimal one. In the former paper, the authors discovered that the number of PCGLS iteration for the optimal preconditioner is $O(\log^2 n)$, while, in the second paper (presented at the Second International Conference on Structured Matrices, at the Hong Kong Baptist University in 2006), the authors stated that the superoptimal one requires $O(\log^3 n)$ iterations. Furthermore, the optimal approach is here the best when considering both time and reconstruction quality.

3.2.2. TEST FH2

In the second test FH2, more regularization capabilities are required since the ill-conditioning becomes more significant. Hence the superoptimal preconditioner usually overcomes the natural and optimal ones which do not possess any regularization feature. Here we show three tables with the best relative restoration error (RRE) $\|x_k - x\|/\|x\|$ and the corresponding iteration k for all the considered instances: Table 4 for the lowest pole degree ($d = 0.37$), Table 5 for the middle pole degree ($d = 0.49$), and Table 6 for the highest pole degree ($d = 0.499$).

We can discuss a wide range of interesting results. In the first case of Table 4, where $d = 0.37$, the first column related to the noiseless data shows that the optimal preconditioner achieves the best performances among all. In particular, although the Relative Restoration Errors by the optimal preconditioners are similar to the other instances (non-preconditioned case, natural and superoptimal), the number of iterations is substantially lower when compared with both the non-preconditioned and the superoptimal preconditioned cases: for instance, the PCGLS with optimal preconditioner requires 18 iterations, while the (non-preconditioned) CGLS and the PCGLS with

Table 4
Best relative restoration errors (RREs) and number of iterations for TEST FH2 with $d = 0.37$

	TEST FH2: $d = 0.37$						
Noise	0	5.00e−07	1.00e−06	5.00e−06	1.00e−05	5.00e−05	1.00e−04
$n = 256$							
No prec	7.1528e−12	2.7676e−03	4.9872e−03	1.8070e−02	3.0625e−02	9.4835e−02	1.4954e−01
	130	22	20	15	14	10	9
Nat	1.2041e−11	6.1038e−03	1.2204e−02	6.1040e−02	1.2209e−01	6.1049e−01	1.2210e00
	22	7	7	7	7	7	7
Opt	9.9348e−12	6.1047e−03	1.2210e−02	6.0908e−02	1.2119e−01	6.0539e−01	1.2105e00
	14	7	7	4	4	4	2
Super	6.5797e−12	4.9774e−03	8.4261e−03	2.9111e−02	4.9111e−02	1.6658e−01	2.4996e−01
	116	14	11	8	6	4	4
$n = 512$							
No prec	2.9969e−11	3.4769e−03	6.1271e−03	2.1351e−02	3.5852e−02	1.0750e−01	1.6577e−01
	254	21	19	15	14	10	9
Nat	5.4207e−11	1.8422e−02	3.6843e−02	1.8424e−01	3.6848e−01	1.8424e00	3.6849e00
	20	8	7	7	7	7	7
Opt	5.6312e−11	1.8423e−02	3.6847e−02	1.8424e−01	3.6627e−01	1.8011e00	3.5236e00
	16	7	7	7	4	2	2
Super	2.5997e−11	8.6365e−03	1.3675e−02	4.4101e−02	7.5190e−02	2.1353e−01	3.3775e−01
	106	11	9	7	6	4	3
$n = 1024$							
No prec	1.1939e−10	3.7670e−03	6.5350e−03	2.2692e−02	3.7589e−02	1.0939e−01	1.6911e−01
	396	20	19	15	14	10	9
Nat	2.2648e−10	5.1272e−02	1.0256e−01	5.1286e−01	1.0258e00	5.1171e00	1.0172e01
	18	7	7	7	7	2	2
Opt	2.2384e−10	5.1287e−02	1.0257e−01	5.1237e−01	1.0246e00	4.9638e00	9.8815e00
	15	8	8	5	5	2	2
Super	1.0863e−10	1.2054e−02	1.8888e−02	5.8924e−02	9.5796e−02	2.7878e−01	3.8755e−01
	137	10	9	6	5	4	3

superoptimal preconditioner require 130 and 118 iterations, respectively, for $n = 256$. However, the noisy cases of the same Table 4 reveal a different behavior, and therein the superoptimal preconditioner leads to the best restorations. Indeed, the superoptimal approach is the only one that gives the same level, or at most a bit worse, of accuracy of the non-preconditioned case, due to its natural regularization capabilities. Optimal and natural ones give rise to lower quality in the restorations, and the related values of the RREs are one order of magnitude worse than the superoptimal (see, for instance, the column with 1.00e−05 of relative noise, for $n = 512$, where: CGLS RRE = 0.035 in 14 iterations, PCGLS-Natural RRE = 0.36 in 7 iterations, PCGLS-Optimal RRE = 0.36 in 4 iterations, PCGLS-Superoptimal RRE = 0.075 in 6 iterations).

The regularization capabilities of the superoptimal preconditioner can be observed more clearly in Fig. 3, where the best restorations related to the second-last column of Table 4 (with relative noise 5.00e−05) are shown. The corresponding convergence history for $n = 1024$ is reported in Fig. 6(a).

Regarding the case of Table 5 ($d = 0.49$), where the ill-conditioning is higher than Table 4 ($d = 0.37$), the filtering capabilities of the superoptimal preconditioner leads to restorations with relative restoration errors similar to (or, in some instances, even a bit lower than) the non-preconditioned case, within about the same number of iterations. Here to be short we give only

Table 5
Best relative restoration errors and number of iterations for TEST FH2 with $d = 0.49$

				TEST FH2: $d = 0.49$, $n = 1024$			
Noise	0	1.00e−09	5.00e−09	1.00e−08	5.00e−08	1.00e−07	5.00e−07
No prec	6.5234e−07	1.0545e−02	3.3539e−02	5.3222e−02	1.4280e−01	2.0180e−01	4.3286e−01
	452	31	23	21	16	12	9
Nat	1.9598e−05	5.8109e−01	2.8911e00	5.6189e00	2.5087e01	4.7991e01	2.5425e02
	37	17	8	8	6	6	6
Opt	1.7254e−05	5.8092e−01	2.9020e00	4.8711e00	1.8363e01	3.6232e01	1.8019e02
	17	6	6	2	2	2	2
Super	4.3690e−06	1.1907e−02	3.7300e−02	5.8994e−02	1.5511e−01	2.1980e−01	4.3776e−01
	500	32	24	20	12	10	6

Table 6
Best relative restoration errors and number of iterations for TEST FH2 with $d = 0.499$

				TEST FH2: $d = 0.499$, $n = 1024$			
Noise	0	5.00e−10	1.00e−09	5.00e−09	1.00e−08	5.00e−08	1.00e−07
No prec	1.1124e−05	1.4885e−01	2.0998e−01	4.5572e−01	5.3767e−01	7.6316e−01	8.7839e−01
	261	18	13	9	8	4	3
Nat	1.7242e−04	4.5391e00	9.2637e00	3.6624e01	7.5006e01	3.9079e02	7.8115e02
	500	51	51	30	30	43	42
Opt	9.9987e−05	3.9114e00	7.2373e00	3.5183e01	7.0297e01	3.5136e02	7.0272e02
	25	2	2	2	2	2	2
Super	4.6949e−04	5.2336e−02	8.1272e−02	1.8722e−01	2.7116e−01	4.4811e−01	5.3473e−01
	500	32	24	15	10	5	4

the results related to the size $n = 1024$, since for the two other sizes $n = 256$, $n = 512$ the behavior is about the same. As already pointed out in the previous case, natural and optimal preconditioners cannot be used when the noise increases (see the last columns, where the RREs are often much higher than 100%). The superoptimal preconditioner requires more iterations than the other two preconditioners, but this slower convergence speed guarantees noise filtering and numerical stability. As instance, we can consider the second column of the Table 5 with 1.00e−09 of relative noise, where: CGLS RRE = 0.010 in 31 iterations, PCGLS-Natural RRE = 0.581 in 17 iterations, PCGLS-Optimal RRE = 0.581 in 6 iterations, PCGLS-Superoptimal RRE = 0.011 in 32 iterations. Fig. 4 shows the corresponding restorations: we can observe that the restoration of the superoptimal preconditioner is really very good, and much more accurate than the natural and optimal ones. The related convergence history of Fig. 6(b) confirms the equivalence between the CGLS and the superoptimal PCGLS, both much more effective than the other preconditioners. Hence test FH2 with $d = 0.49$ of Table 5 shows that the PCGLS with the superoptimal preconditioner can be used with higher safety since often, although it does not improve the results substantially, does not degrade the performances of the (non-preconditioned) CGLS.

In Table 6, the results related to the case $d = 0.499$ are reported (again for shortness we plot results for $n = 1024$ only), where the ill-conditioning is even larger. This case with $d = 0.499$ is the most interesting one, since the ill-position of the problem is so high that the regularization features of the non-preconditioned method CGLS are not sufficient to provide good restorations. Here the preconditioned method PCGLS with superoptimal preconditioner even excels the

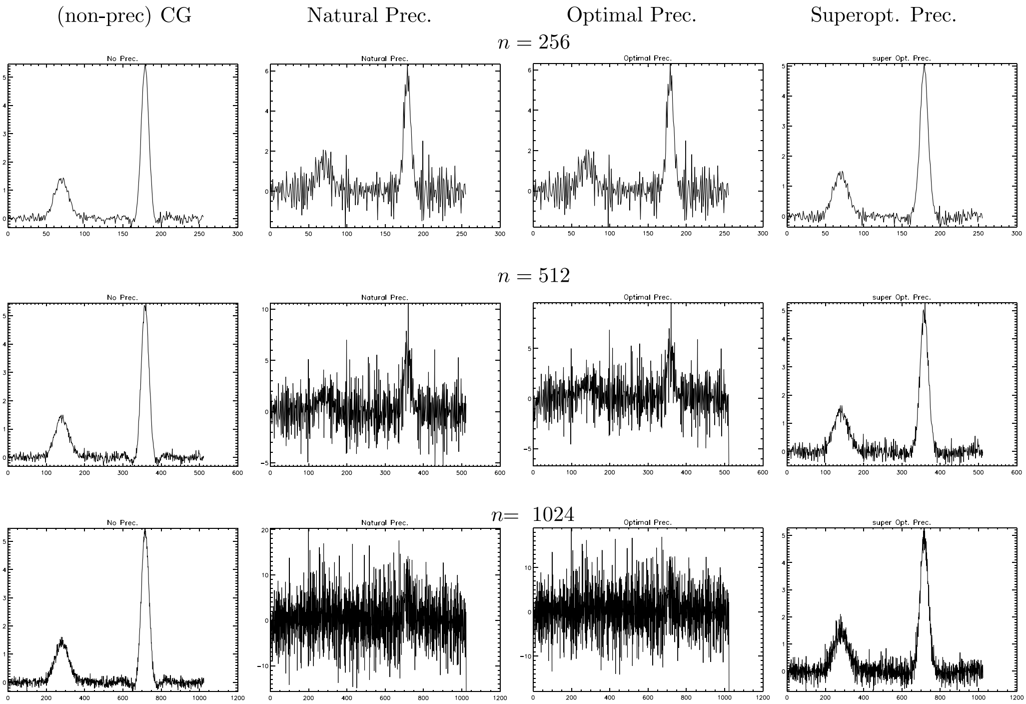

Fig. 3. Best restorations for TEST FH2 with $d = 0.37$ and relative noise 5.00e−05 (see the numerical values in the second-last column of Table 4).

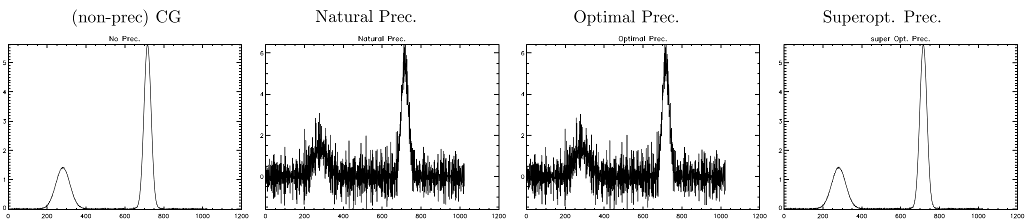

Fig. 4. Best restorations for TEST FH2 with $d = 0.49$ and relative noise 1.00e−09, for $n = 1024$ (see the numerical values in the second column of Table 5).

non-preconditioned CGLS. In particular, here the superoptimal preconditioner slows down very much the convergence speed, and the noisy components can be controlled and suitably filtered out. In addition, the choice of the stopping iteration, which is well known to be a so crucial and difficult task, is simplified, since the restoration errors – which are always unknown in real applications – stay small for several iterations before starting to diverge (see the related convergence history of Fig. 6(c) where the curve of the RREs by superoptimal is quite flat around the minimal value).

Notice that, except the trivial case with no noise, here natural and optimal preconditioners give completely wrong restorations (RREs generally much greater than 100%!). As an example, we consider the second column of Table 6 with 5.00e−10 of relative noise, where CGLS RRE = 0.14 in 18 iterations, PCGLS-Natural RRE = 4.5 in 51 iterations, PCGLS-Optimal RRE = 3.9 in 2 iterations, PCGLS-Superoptimal RRE = 0.052 in 32 iterations. We remark that this latter value

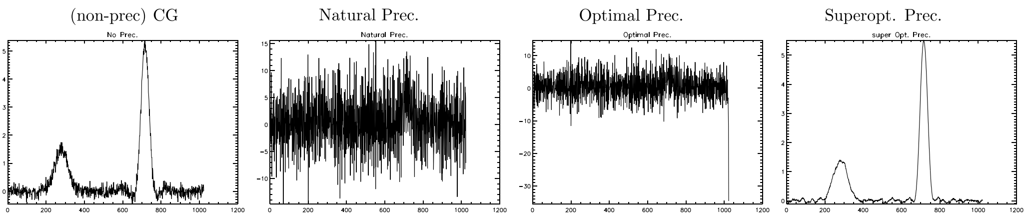

Fig. 5. Best restorations for TEST FH2 with $d = 0.499$ and relative noise 5.00e−10, for $n = 1024$ (see the numerical values in the second column of Table 6).

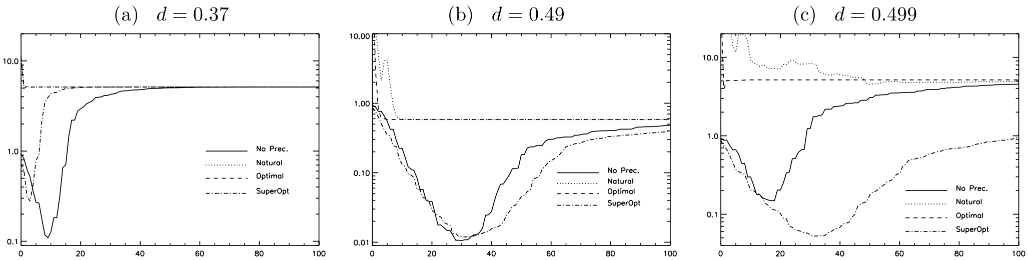

Fig. 6. RREs versus Iterations for TEST FH2 within the first 100 iterations, $n = 1024$, for $d = 0.37$, $d = 0.49$ and $d = 0.499$, with corresponding relative noise on input data 5.00e−5, 1.00e−9 and 5e−10.

of RRE (0.052) is much lower than the one of the non-preconditioned case (0.14). Fig. 5 shows the corresponding restorations, and Fig. 6(c) the convergence history for $n = 1024$. As already mentioned, the right iterations where to stop the algorithm can be chosen more easily (all the iterations between the 15th and the 50th give good restorations, which are all better than the best one, i.e. the 18th, of the non-preconditioned case). It is worth noting that here the superoptimal preconditioning is the only strategy that guarantees a good level of noise filtering, among all the considered ones.

3.2.3. TEST FH3

In the last test FH3, the ill-conditioning is extremely high. In these case, although the natural and optimal give totally unsatisfactory results even for low noise (see all but the first two columns of Table 7, with $n = 1024$ and $d = 0.49$, where the RREs are always greater than 50% for all those very low noise levels), the superoptimal preconditioner leads to a number of iterations greater than the non-preconditioned case, with a very similar restoration accuracy. In these cases, the regularization features of the superoptimal preconditioner are not sufficient, and it can be useful to apply the already mentioned filtered extensions [18] which are based on the superoptimal idea.

3.3. Signal restoration versus Fisher–Hartwig singularities

In [15,12], when considering signal/image restoration problems the following observations were proposed:

- The superoptimal approximation leads to a basic level of regularization without further operations: for instance, in some cases like Large Binocular Telescope (LBT) applications the superoptimal preconditioning is enough for obtaining qualitatively good images. This feature

Table 7
Best relative restoration errors and number of iterations for TEST FH3

	TEST FH3: $d = 0.49$, $n = 1024$						
Noise	0	1.00e−10	5.00e−10	1.00e−09	5.00e−09	1.00e−08	5.00e−08
No prec	3.0076e−04	3.5780e−01	5.3663e−01	6.2103e−01	7.9395e−01	8.4407e−01	9.2292e−01
	214	19	12	9	6	5	3
Nat	1.9257e−01	3.6050e01	1.5947e02	4.4474e02	1.0649e03	1.3634e03	5.0302e03
	134	8	8	8	1	1	1
Opt	1.0952e00	1.4353e01	5.7366e01	1.1357e02	5.6551e02	1.1308e03	5.5533e03
	8	2	2	2	2	2	1
Super-Opt	2.1875e−03	1.6852e−01	3.2411e−01	3.9094e−01	5.6414e−01	5.9559e−01	8.3159e−01
	500	60	30	24	11	8	4

is not common in other techniques where the preconditioner has to be filtered in some way and this operation is often very difficult to do (see e.g. [20]).

- The level of regularization achieved by the superoptimal operator can be not sufficient but it is higher when compared with any other classical preconditioning operator (natural, optimal, etc. [8]); therefore, the superoptimal approximation can be used as starting point and can be favorably combined with Tikhonov like shift (see e.g. [2,17]) of the spectrum (as in [26]), with the stabilization idea used e.g. in [27], or with the filtering idea used e.g. in [18].
- The computational cost of any iteration is $\mathrm{O}(N \log N)$ which is the minimal one when dealing with trigonometric algebras.

Here we want to explain why the latter signal restoration problems are equivalent, by a spectral point of view, to the case of Toeplitz systems with Fisher–Hartwig singularities of type (13) considered here. By virtue of this equivalence, the above conclusions can be applied to unbounded symbols (13) too. In addition, since the case of images is similar to signals, being a simple two-level generalization, the same conclusions can be also extended to the corresponding two-dimensional versions.

Consider the matrix $T_n = \widetilde{T}_n(h), h = f^c, c = 1, 2, 3,$ and f as in (13). Consider λ_n the maximal eigenvalue of T_n (i.e. the spectral norm since T_n is symmetric positive definite) and take the "normalized" version

$$\widehat{T}_n = T_n/\lambda_n.$$

Then the following facts hold true:

1. Any linear system with coefficient matrix $\widehat{T}_n$ is equivalent to a linear system with coefficient matrix T_n (the two systems only differ from the scaling λ_n in the right-hand side).
2. The Strang and the optimal preconditioners of T_n is λ_n times the ones of $\widehat{T}_n$: thanks to (5) and (6), the same is true for the superoptimal preconditioner since the scalar value λ_n is positive.
3. As a consequence of the previous two items, the preconditioned matrices of $\widehat{T}_n$ with the three considered preconditioners exactly coincide with those of T_n.
4. The value λ_n is of the order of n^s, $s = 2cd$, and the pole of function f is located for $\theta = 0$: therefore, in the low frequencies the eigenvalues of $\widehat{T}_n$ are close to 1 and in the middle and high frequencies the related eigenvalues are of the order of n^{-s}, $s = 2cd$.
5. From the third and the fourth items, recalling the classical Gaussian blur for signal and image restoration problems (see e.g. [2,17]), we deduce that the two problems, i.e. signal/image

restoration and Fisher–Hartwig problems, are spectrally equivalent and therefore, it is not surprising that in the above subsections we have observed results similar to those of [15,12].

Finally, we just remark a further similarity between the two problems since in both the cases the coefficients of the Toeplitz matrix (see (14)) are nonnegative (typical of a blurring operation).

Conclusions

In summary, we can state that the superoptimal preconditioner can work as a basic regularizing tool, which, in absence of critical conditions (e.g. high noise level and/or excessive ill-conditioning), provides reconstructions of acceptable quality with fewer iterations or better accuracy with more iterations, when compared to the non-preconditioned method. The advantage becomes much more evident when comparing with the Strang or the optimal preconditioners which provide bad results in presence of noise (even in small percentages when the conditioning is very high). On the other hand, according to the convergence analysis of [23,44], when the solution is in the classical sense i.e. without noise, the superoptimal preconditioner is slightly worse than the other two approaches when the symbol is positive and bounded, also taking into account that the computational cost for its construction is higher (though of the same order). In particular, in this case with positive and bounded symbol, the natural and the optimal approaches are equivalent for moderate ill-conditioning while the optimal preconditioning is superior when the noise is absent and the ill-conditioning is more severe. Furthermore, the difference between the superoptimal and the other two preconditioners increases when dealing with positive, unbounded symbols while it becomes unacceptably large when the symbol has zeros: in the latter case, the reason is due to the rank obstruction argument and in the former case the reason is the type of approximation i.e. the relative approximation (or superoptimal approximation), which is worse than the absolute approximation (or optimal approximation) for large values of the spectrum. It is worth noting that the regularization features of the superoptimal preconditioning can be favorable to improve the reconstructions especially in both of these two last cases.

References

[1] M. Benzi, C. Meyer, M. Tuma, A sparse approximate inverse preconditioner for the conjugate gradient method, SIAM J. Sci. Comput., 17-5 (1996) 1135–1149.
[2] M. Bertero, P. Boccacci, Introduction to Inverse Problems in Imaging, Institute of Physics Publ, Bristol, 1998.
[3] R. Bhatia, Fourier Series, AMS, Providence, 2004.
[4] A. Böttcher, B. Silbermann, Analysis of Toeplitz Operators, Springer-Verlag, Berlin, 1990.
[5] P. Brockwell, R. Davis, Time Series: Theory and Methods, Springer, New York, NY, 1991.
[6] D. Calvetti, L. Reichel, On the regularizing properties of the GMRES method, Numer. Math. 91-4 (2002) 605–625.
[7] R. Chan, X. Jin, M. Yeung, The spectra of super-optimal circulant preconditioned Toeplitz systems, SIAM J. Numer. Anal. 28 (1991) 871–879.
[8] R. Chan, M. Ng, Conjugate gradient methods for Toeplitz systems, SIAM Rev. 38 (1996) 427–482.
[9] R. Chan, P. Tang, Fast band-Toeplitz preconditioners for Hermitian Toeplitz systems, SIAM J. Sci. Comput. 15 (1994) 164–171.
[10] T. Chan, An optimal circulant preconditioner for Toeplitz systems, SIAM J. Sci. Stat. Comput. 9 (1988) 766–771.
[11] F. Costabile, S. Serra-Capizzano, Komleva type expansions and asymptotics for linear operators, Comput. Math. Appl. 43 (2002) 799–820.
[12] F. Di Benedetto, C. Estatico, S. Serra-Capizzano, Superoptimal preconditioned conjugate gradient iteration for image deblurring, SIAM J. Sci. Comput. 26-3 (2005) 1012–1035.
[13] F. Di Benedetto, S. Serra-Capizzano, A unifying approach to abstract matrix algebra preconditioning, Numer. Math. 82-1 (1999) 57–90.

[14] F. Di Benedetto, S. Serra-Capizzano, Optimal multilevel matrix algebra operators, Linear Multilinear Algebra 48 (2000) 35–66.
[15] F. Di Benedetto, S. Serra-Capizzano, A note on the superoptimal matrix algebra operators, Linear Multilinear Algebra 50 (2002) 343–372.
[16] L. Eldén, An algorithm for the regularization of ill-conditioned, banded least squares problems, SIAM J. Sci. Stat. Comput. 5 (1984) 237–254.
[17] H. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
[18] C. Estatico, A class of filtering superoptimal preconditioners for highly ill-conditioned linear systems, BIT 42 (2002) 753–778.
[19] C. Estatico, A classification scheme for regularizing preconditioners, with application to Toeplitz systems, Linear Algebra Appl. 397 (2005) 107–131.
[20] C. Estatico, Regularized fast multiple-image deconvolution for LBT, J. Comput. Appl. Math. 198 (2007) 443–459.
[21] A. Greenbaum, Iterative Methods for Solving Linear Systems, SIAM Publ., Philadelphia, 1997.
[22] T. Huckle, Efficient computation of sparse approximate inverses, Numer. Linear Algebra Appl. 5-1 (1998) 57–71.
[23] C. Hurvich, Y. Lu, On the complexity of the preconditioned conjugate gradient algorithm for solving Toeplitz systems with a Fisher–Hartwig singularity, SIAM J. Matrix Anal. Appl. 27 (2005) 638–653.
[24] D. Jackson, The Theory of Approximation, American Mathematical Society, New York, 1930.
[25] T. Kailath, V. Olshevsky, Displacement structure approach to discrete-trigonometric-transform based preconditioners of G. Strang type and T. Chan type, in: Proceedings of the Workshop on Toeplitz matrices, Cortona, Italy, September 1996; Calcolo, 33, 1996, pp. 191–208.
[26] J. Kamm, J. Nagy, Kronecker product and SVD approximation in image restoration, Linear Algebra Appl. 284 (1998) 177–192.
[27] M. Kilmer, Cauchy-like preconditioners for two-dimensional ill-posed problems, SIAM J. Matrix Anal. Appl. 20 (1999) 777–799.
[28] T. Manteuffel, S. Parter, Preconditioning and boundary conditions, SIAM J. Numer. Anal. 27-3 (1990) 656–694.
[29] D. Noutsos, S. Serra-Capizzano, P. Vassalos, Matrix algebra preconditioners for multilevel Toeplitz systems do not insure optimal convergence rate, Theoret. Comput. Sci. 315 (2004) 557–579.
[30] Y. Saad, Iterative Methods for Sparse Linear Systems, PWS Publishing, Boston, MA, 1996.
[31] S. Serra-Capizzano, Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems, Math. Comp. 66 (1997) 651–665.
[32] S. Serra-Capizzano, On the extreme eigenvalues of Hermitian (block) Toeplitz matrices, Linear Algebra Appl. 270 (1998) 109–129.
[33] S. Serra-Capizzano, An ergodic theorem for classes of preconditioned matrices, Linear Algebra Appl. 282 (1998) 161–183.
[34] S. Serra-Capizzano, How to choose the best iterative strategy for symmetric Toeplitz systems, SIAM J. Numer. Anal. 36-4 (1999) 1078–1103.
[35] S. Serra-Capizzano, A Korovkin-type theory for finite Toeplitz operators via matrix algebras, Numer. Math. 82-1 (1999) 117–142.
[36] S. Serra-Capizzano, Korovkin tests, approximation, and ergodic theory, Math. Comp. 69 (2000) 1533–1558.
[37] S. Serra-Capizzano, Practical band Toeplitz preconditioning and boundary layer effects, Numer. Algorithms 34 (2003) 427–440.
[38] S. Serra-Capizzano, P. Tilli, On unitarily invariant norms of matrix valued linear positive operators, J. Inequal. Appl. 7-3 (2002) 309–330.
[39] G. Strang, A proposal for Toeplitz matrix calculations, Stud. Appl. Math. 74 (1986) 171–176.
[40] V. Strela, E. Tyrtyshnikov, Which circulant preconditioner is better?, Math. Comp. 65-213 (1996) 137–150.
[41] E. Tyrtyshnikov, Optimal and superoptimal circulant preconditioners, SIAM J. Matrix Anal. Appl. 13 (1992) 459–473.
[42] E. Tyrtyshnikov, Circulant preconditioners with unbounded inverses, Linear Algebra Appl. 216 (1995) 1–23.
[43] E. Tyrtyshnikov, N. Zamarashkin, Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships, Linear Algebra Appl. 270 (1998) 15–27.
[44] S. Vong, W. Wang, X. Jin, A note on the complexity of the PCG algorithm for solving Toeplitz systems with a Fisher–Hartwig singularity, in: Z. Yao, M. Yuan, Y. Chen (Eds.), Computational Methods in Engineering and Science, Proceedings of EPMESC X, Tsinghua University Press and Springer, 2006, pp. 903–907.
[45] A. Zygmund, Trigonometric Series, Cambridge University Press, Cambridge, 1959.